WHAT IS CLAIMED IS:

1. A method of determining the statistical significance of a difference between haplotype frequency profiles of at least two groups of individuals comprising:

determining the combined likelihood that said at least two groups of individuals are derived from the same distribution of haplotypes;

determining the sum of the separate likelihoods that each of said at least two groups of individuals are derived from the same distribution of haplotypes; determining the difference of said sum and said combined likelihood; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

2. The method of claim 1, further comprising calculating all possible single-haplotype chi-square tests prior to determining the significance of the difference between said sum and said combined likelihood.

3. The method of claim 1, further comprising assessing the statistical significance of individual haplotypes using an odds ratio or a P-excess value.

4. A system for determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

first instructions for determining the combined likelihood that said at least two groups of individuals are derived from the same distribution of haplotypes;

second instructions for determining the sum of the separate likelihoods that each of said at least two groups of individuals are derived from the same distribution of haplotypes;

third instructions for determining the difference of said sum and said combined likelihood; and

fourth instructions for determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

5. The system of claim 4, further comprising fifth instructions for calculating all possible single-haplotype chi-square tests prior to determining the significance of the difference between said sum and said combined likelihood.

6. The system of claim 4, further comprising fifth instructions for assessing the statistical significance of individual haplotypes using an odds ratio or a P-excess value.

7. A programmed storage device comprising instructions that when executed perform a method comprising:

determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals by comparing the final likelihood that all groups of individuals come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

8. The programmed storage device of claim 7, further comprising instructions that when executed perform a method of calculating all possible single-haplotype chi-square tests prior to determining the significance of the difference between said sum and said combined likelihood.

9. The programmed storage device of claim 7, further comprising instructions that when executed perform a method of assessing the statistical significance of individual haplotypes using an odds ratio or a P-excess value.

10. A method of estimating haplotype frequencies for single nucleotide polymorphisms in groups of individuals comprising:

estimating all haplotype and diplotype probabilities for said groups of individuals using an estimation-maximization process;

storing said probabilities; and

repeating said estimation-maximization process using random starting values.

11. The method of claim 10, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing said estimations.

12. A computer system for estimating haplotype frequencies for single nucleotide polymorphisms in groups of individuals comprising:

first instructions that when executed perform a method of estimating all haplotype and diplotype probabilities using an estimation-maximization process;

second instructions that when executed perform a method of storing said haplotype and diplotype probabilities; and

third instructions that when executed perform said estimation-maximization process that is automatically repeated using random starting values.

13. The computer system of claim 12, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing estimations.

14. A programmed storage device comprising estimation-maximization instructions that when executed perform the method of:

estimating haplotype frequencies for single nucleotide polymorphisms in groups of individuals comprising estimating and storing all haplotype and diplotype probabilities using an estimation-maximization process; and

repeating said estimation-maximization process using random starting values.

15. The programmed storage device of claim 14, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing estimations.

16. A method of determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

estimating haplotype frequencies using single nucleotide polymorphism data for each group individually and for each group in combination with another group, wherein all haplotype and diplotype probabilities are calculated once and then stored, and wherein a maximization process is automatically repeated for each group using random starting values in order to determine final likelihoods;

comparing the final likelihood that all groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately to determine their difference; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

17. The method of claim 16, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing operations.

18. A system for determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

a first module configured to estimate haplotype frequencies using single nucleotide polymorphism data for each group individually and for each group in combination with another group, wherein all haplotype and diplotype probabilities are calculated once and then stored, and wherein the maximization process is automatically repeated for each group using random starting values, to determine final likelihoods;

a second module configured to compare the final likelihood that all groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately to determine their difference; and

a third module configured to determine the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

19. The system of claim 18, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing estimations.

20. A programmed storage device comprising instructions that when executed perform a method of determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising

a first module adapted to perform a method of estimating haplotype frequencies using single nucleotide polymorphism data for each group individually and for each group in combination with the other group, wherein all haplotype and diplotype probabilities are calculated once and then stored, and wherein the maximization process is automatically repeated for each group using random starting values to determine final likelihoods;

a second module adapted to compare the final likelihood that all groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately to determine their difference; and

a third module adapted to determine the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

21. The programmed device of claim 20, wherein all haplotypes are coded with binary mask arrays, and wherein identical genotypes are grouped prior to performing estimations.

22. A method of determining an association between a haplotype and a phenotype, comprising:

estimating haplotype frequencies using single nucleotide polymorphism data for an affected group and an unaffected group individually and in combination with another group, wherein all haplotype and diplotype probabilities are calculated once and then stored, and wherein a maximization process is automatically repeated for each group using random starting values to determine final likelihoods;

comparing the final likelihood that both groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately to determine their difference; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and determine whether a statistically significant association exists between said haplotype and said phenotype.

23. A method of determining an association between a haplotype and a phenotype, comprising:

estimating haplotype frequencies using single nucleotide polymorphism data for an affected group and an unaffected group individually and in combination with another group, wherein all haplotype and diplotype probabilities are calculated once;

storing said probabilities; and

repeating a maximization process for each group using random starting values to determine whether a statistically significant association exists between said haplotype and said phenotype.

24. A method of detecting an association between a haplotype and a phenotype, comprising:

comparing a final likelihood that members of an affected group and an unaffected group come from the same distribution of haplotypes with the sum of the final likelihoods for each of said groups separately to determine their difference; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and whether a statistically significant association exists between said haplotype and said phenotype.

25. A system for detecting an association between a haplotype and a phenotype, comprising:

first instructions for estimating haplotype frequencies using single nucleotide polymorphism data for an affected group and an unaffected group individually and in combination, wherein all haplotype and diplotype probabilities are calculated once, and wherein the maximization process is automatically repeated using random starting values to determine final likelihoods;

second instructions for comparing the final likelihood that both groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

third instructions for determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and determine whether a statistically significant association exists between said haplotype and said phenotype.

26. A system for detecting an association between a haplotype and a phenotype, comprising:

instructions for estimating haplotype frequencies using single nucleotide polymorphism data for an affected and an unaffected group individually and in combination, wherein all haplotype and diplotype probabilities are calculated once; and

repeating a maximization process using random starting values to determine whether a statistically significant association exists between said haplotype and said phenotype.

27. A system for detecting an association between a haplotype and a phenotype, comprising:

first instructions for comparing the final likelihood that the members of an affected and an unaffected group come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately;

second instructions for determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and whether a statistically significant association exists between said haplotype and said phenotype.

28. A programmed storage device comprising instructions that when executed perform a method of detecting an association between a haplotype and a phenotype, comprising:

estimating haplotype frequencies using single nucleotide polymorphism data for an affected and an unaffected group individually and in combination, wherein all haplotype and diplotype probabilities are calculated once and are stored, and wherein the maximization process is automatically repeated using random starting values to determine final likelihoods;

comparing the final likelihood that both groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and determine whether a statistically significant association exists between said haplotype and said phenotype.

29. A programmed storage device comprising instructions that when executed perform a method of detecting an association between a haplotype and a phenotype, comprising:

estimating haplotype frequencies using single nucleotide polymorphism data for an affected and an unaffected group individually and in combination, wherein all haplotype and diplotype probabilities are calculated once; and

repeating a maximization process using random starting values to determine whether a statistically significant association exists between said haplotype and said phenotype.

30. A programmed storage device comprising instructions that when executed perform a method of detecting an association between a haplotype and a phenotype, comprising:

comparing a likelihood that members of an affected group and an unaffected group come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately;

determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes; and

determining whether a statistically significant association exists between said haplotype and said phenotype.

31. A computer-readable data signal embedded in a transmission medium that when executed performs a method of determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

code segments comparing the final likelihood that all groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

code segments determining the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

32. A wide area computer network for determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

a server comprising single nucleotide polymorphism data; and

a workstation comprising instructions for estimating haplotype frequencies using said nucleotide polymorphism data for each group individually and in combination with the other group, wherein all haplotype and diplotype probabilities are calculated once and are stored, and wherein the maximization process is automatically repeated using random starting values.

33. The wide area computer network of claim 32, wherein said network comprises the Internet.

34. The wide area computer network of claim 32, wherein said instructions are stored in a memory.

35. The wide area computer network of claim 32, wherein said instructions are stored in a code segment.

36. A computer-readable data signal embedded in a transmission medium that when interpreted performs a method determining the statistical significance of the difference between haplotype frequency profiles of at least two groups of individuals, comprising:

first signals adapted to perform a method of estimating haplotype frequencies using single nucleotide polymorphism data for each group individually and in combination with the other group, wherein all haplotype and diplotype probabilities are calculated once and are stored, and wherein a maximization process is automatically repeated using random starting values, to determine final likelihoods;

second signals adapted to compare the final likelihood that all groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

third signals adapted to determine the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes.

37. A computer system for detecting an association between a haplotype and a phenotype, comprising:

a first code segment configured to estimate haplotype frequencies using single nucleotide polymorphism data for an affected and an unaffected group individually and in

combination, wherein all haplotype and diplotype probabilities are calculated once and are stored, and wherein a maximization process is automatically repeated using random starting values to determine final likelihoods;

a second code segment configured to compare the final likelihood that both groups come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately; and

a third code segment configured to determine the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and determine whether a statistically significant association exists between said haplotype and said phenotype.

38. A computer-readable data signal embedded in a transmission medium that when executed performs a method of detecting an association between a haplotype and a phenotype, comprising:

a first signal for estimating haplotype frequencies using single nucleotide polymorphism data for an affected and an unaffected group individually and in combination, wherein all haplotype and diplotype probabilities are calculated once and are stored; and

a second signal for repeating a maximization process using random starting values to determine whether a statistically significant association exists between said haplotype and said phenotype.

39. A wide area computer system for detecting an association between a haplotype and a phenotype, comprising:

a first memory comprising first code segments adapted to compare the final likelihood that the members of an affected and an unaffected group come from the same distribution of haplotypes with the sum of the final likelihoods for each group separately;

a second memory comprising second code segments adapted to determine the significance of this difference by simulating hypothetical groups by randomly permuting the haplotypes between groups to determine the probability that the groups do not come from the same distribution of haplotypes and whether a statistically significant association exists between said haplotype and said phenotype.